GLOBAL BIOGEOCHEMICAL CYCLES, VOL. 16, NO. 1, 1014, 10.1029/2000GB001381, 2002

Upscaling methane emissions from rice paddies: Problems and possibilities

P. M. van Bodegom[1]
Laboratory of Microbiology, Wageningen University, Wageningen, Netherlands
Department of Theoretical Production Ecology, Wageningen University, Wageningen, Netherlands

Peter H. Verburg and Hugo A. C. Denier van der Gon
Laboratory of Soil Science and Geology, Wageningen University, Wageningen, Netherlands

Received 4 December 2000; revised 17 July 2001; accepted 27 August 2001; published 6 March 2002.

[1] Global methane emission estimates depend highly on the models, techniques, and databases used. Since emissions cannot be measured directly at large scales, it is impossible to judge which estimate is more realistic. In this paper, different aspects of uncertainty in upscaling methane emissions from rice paddies are discussed. These aspects are visualized by a case study on the spatial upscaling of methane emissions from the island of Java, Indonesia. The first aspect concerns process information. An approach to incorporate this information in a simplified but process-based way in predictive models is discussed. Sources of uncertainty include the methane emissions measurements, processes quantification, process simplification, and the use of data transfer functions. Data availability of input parameters, the second aspect, is uncertain because of differences between different data sources, the use of data sources for purposes not originally planned for, and the scale at which data are available. Data interpolation in combination with nonlinear model responses introduces scaling errors, the third aspect. Data accuracy introduced the highest uncertainties in emission estimates but is rarely accounted for in the estimation of global emissions. *INDEX TERMS:* 1615 Global Change: Biogeochemical processes (4805); 3322 Meteorology and Atmospheric Dynamics: Land/atmosphere interactions; 3210 Mathematical Geophysics: Modeling; 6309 Policy Sciences: Decision making under uncertainty; *KEYWORDS:* spatial variability, model, uncertainty, scale, interpolation

1. Introduction

[2] Methane (CH_4) emissions are an important greenhouse gas, accounting for 15–20% of the radiative forcing added to the atmosphere [*Houghton et al.*, 1996]. Total global CH_4 emissions are estimated rather accurately at 600 ± 80 Tg yr^{-1} [*Lelieveld et al.*, 1998] on the basis of the assumption that the total source must balance the global CH_4 sinks plus annual atmospheric CH_4 increase. However, source-specific emission estimates are difficult to obtain and depend on the models, techniques, and databases used. Source-specific methane emissions can be estimated by downscaling using atmospheric chemistry-transport models, isotopic signatures, and observational data. A recent estimate for methane emissions from rice paddies using this approach is 80 ± 50 Tg yr^{-1} [*Lelieveld et al.*, 1998]. Alternatively, attempts can be made to understand and explain the complex system by integrating and aggregating spatial information on methane emissions in higher-scale estimates: upscaling [see, e.g., *King*, 1991; *Rosswall et al.*, 1988]. Upscaling is often the only option to derive regional source-specific emission estimates and can be carried out in many different ways [*Denier van der Gon et al.*, 2000]. Efforts of upscaling have resulted in different global estimates of methane emissions from rice paddies, varying from 117 ± 50 Tg yr^{-1} [*Holzapfel-Pschorn and Seiler*, 1986] to 50 ± 20 Tg yr^{-1} [*Neue*, 1997]. Since emissions cannot yet be measured accurately at a national or global scale, it is difficult, if not impossible, to judge which estimate is more realistic.

[3] One principal reason for the large uncertainties in global estimates is the large intrinsic spatial and temporal (diurnal and seasonal) variability in methane emissions [e.g., *Denier van der Gon and Neue*, 1995; *Nouchi et al.*, 1994]. This large variability is the result of complex nonlinear interactions between underlying processes, i.e., methane production, methane oxidation, and methane transport. These interactions are only partly understood, and the resulting variation in methane emissions is not fully explained by correlations with environmental variables [*Walter et al.*, 1996].

[4] Other crucial factors in upscaling are data handling and data inclusion. For example, straw management has a large influence on methane emissions [*Denier van der Gon et al.*, 2000], while information on local practice and variability in these practices throughout the region is difficult to collect and often not included in regular census surveys. In addition, handling of the different scales involved is a crucial factor. Most estimates assume a homogeneous methane emission within each country, while it is known that methane emissions vary widely depending on local conditions and thus that scaling problems may occur.

[5] So a range of different factors influences the accuracy, reliability, and uncertainty in methane emission estimates from rice paddies based on upscaling. The objective of this paper is to discuss the different aspects of accuracy and reliability to be considered in upscaling methane emission estimates from rice paddies. The uncertainties in the various aspects are discussed

[1]Now at Department of Systems Ecology, Faculty of Biology, Free University, Amsterdam, Netherlands.

0886-6236/02/2000GB001381

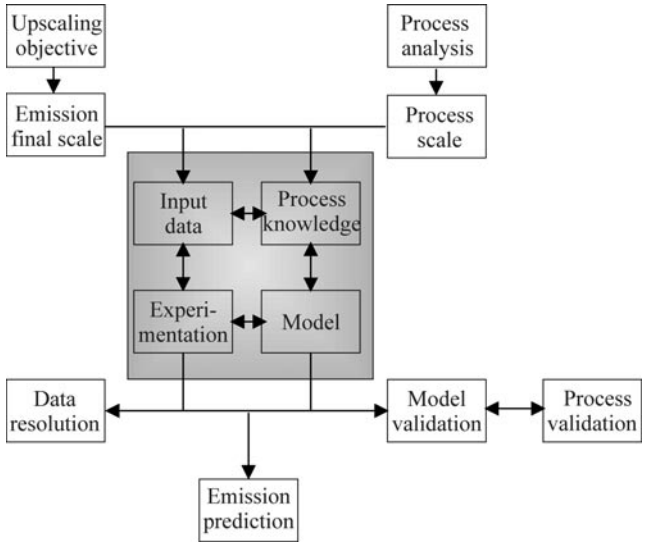

Figure 1. Schematic representation of the analysis applied to come to a model that predicts methane emissions and that can be used for upscaling.

and visualized by a case study on the spatial upscaling of methane emission estimates from the island of Java in Indonesia, focusing on the influence of spatially explicit soil information. The three crucial factors, process information, data, and scale, are treated in separate sections to facilitate quantification, if possible, of the uncertainties caused by these factors.

2. Process Information and Upscaling

2.1. Conditions on Process Information in Upscaling

[6] The processes involved in methane emissions can only be measured up to a scale of $\sim$1 m^2. Methane emissions itself can be measured at a plot scale ($\sim$1 m^2) using a closed chamber technique [e.g., *Wassmann et al.*, 1996] or at a somewhat larger scale, up to a hectare, by micrometeorological techniques [*Simpson et al.*, 1995].

[7] Measurements of CH_4 fluxes at the regional scale are feasible using aircraft [*Choularton et al.*, 1995]. However, this is very expensive, not done in the tropics where most of the rice emissions occur, and more important, not likely to cover a semi-continuous period such as a full rice cropping cycle. This means that emissions at regional scales have to be estimated in a different way, i.e., modeled, and unfortunately, implies that these modeled emissions cannot be validated.

[8] There are important conditions that have to be accounted for in models used for upscaling. Such a model should not require many spatially specific data because only a limited number of variables is easy to obtain in regional databases. This condition poses severe limitations to the level of detail that can be included in a model because more detail generally requires more detailed input data. Furthermore, to apply the model at other sites than those for which the model was validated, coincidental agreement between data must be avoided, and quantification of uncertainties must be allowed. Extrapolation of model results is therefore less risky with process-based models than with empirical models, as similar processes will occur around the world. The central issue is then to quantify all processes that are truly essential to allow such extrapolation.

2.2. Model Development for Upscaling

[9] In the process-based modeling approach a compromise has to be found between the inclusion of driving variables (to describe processes as mechanistically as possible to allow general applicability) and the limited information available on driving variables in databases at regional scales to base such a process description on. This knowledge gap is due to the large difference in process scale and the final scale of the emission estimate because the most important processes leading to methane emissions occur at a very small scale (at the scale of a single plant) and are interdependent. During upscaling the process scale and target scale should remain connected, which is possible only if the key processes and their interactions are accounted for and if these processes are related to available data at the target regional scale. A comprehensive methodology that allows this is depicted in Figure 1.

[10] The approach uses an analysis by modeling and experimentation to distinguish and quantify the processes underlying methane emissions and the driving variables. The individual processes are well described and reviewed by, e.g., *Conrad* [1993] and *Segers* [1998]. On the basis of that information the processes can be linked to available data and integrated to higher-scale outcomes. The relations between individual processes and available data are quantified in a model, securing the link between the scales. While doing this, care should be taken not to combine processes with relative rates differing some orders of magnitude to avoid model instabilities. Moreover, careful choice of system boundaries should be made such that the outside affects the system but the system hardly affects processes outside the boundaries (see section 4.1). This makes the description of the feedbacks and interactions between the processes the crucial step in the approach. Quantitative process knowledge on these interactions is scarce, while it determines, to a large extent, the sensitivity of and variability in processes. To quantify the interactions between the processes leading to methane emissions from rice paddies without increasing the number of input variables, some major simplifications have to be made. The most important simplification made in the model used in this case study was the distinction between a rhizosphere and a bulk soil compartment. All processes have distinct features near and away from the roots [*van Bodegom et al.*, 2001a; *Segers and Rappoldt*, 2000]. The major spatial heterogeneity at the process scale is thus explicitly identified, and the soil is considered to be homoge-

Table 1. Controlling Variables (and Their Values for Variables Other Than Soil Properties) Used to Calculate CH_4 Emissions With the Model of *van Bodegom et al.* [2001a, 2001b]

Input Parameter	Value	Source
Soil organic carbon	spatial explicit values	CSAR or FAO data
Reducible iron content	spatial explicit values	CSAR or FAO data
Rice distribution area	spatial explicit values	*Verburg et al.* [1999]
Rice variety[a]	IR64	*BPS* [1996]
Rice yield	5.1 t ha^{-1}	*BPS* [1995]
NO_{3-} or SO_4-containing fertilizers[b]	34 kg ha^{-1}	*BPS* [1994]
Straw input	30% of the yield	*Neue et al.* [1990]
Length of growing season	119 days	*BPS* [1996]
Temperature	26°C	*Verburg et al.* [1999]
Water management[c]	continuously flooded	

[a] This is the dominant rice variety on Java.
[b] Other inorganic fertilizers, like urea, do not directly influence CH_4 emission. The fertilization effect on rice production is included in the average yield.
[c] A significant part of the rice fields on Java is not continuously flooded (rain-fed rice), but this was not accounted for in this study.

neous within the compartments. With this simplification, process-based functional relationships between the processes and the available data could be established and enable understanding and prediction of the variability found in methane emissions. *van Bodegom et al.* [2001a] described a model made according to this methodology. We use the *van Bodegom et al.* [2001a] model, extended with a description for intermittent drainage and for situations in which rice yield is not optimal [*van Bodegom et al.*, 2000], in our case study. However, other process-based models may be equally suitable for upscaling.

[11] The model needs information on soil reducible iron content and soil organic carbon, rice variety and rice yield, inorganic and organic fertilizer input, length of the growing season, and temperature. In this case study, only soil properties in addition to rice area estimates are treated spatially explicit. The influence of other input parameters is not considered to facilitate analysis of upscaling effects. For input parameters other than soil properties, weighted average values were estimated and used (Table 1).

2.3. Uncertainties in Model Outcomes

[12] A model always remains an approximation of reality and is thus a source of uncertainty. This uncertainty can be attributed to different sources. A first source is the measurement of methane emissions itself. A common procedure is the measurement with closed chambers where methane emission rates are determined from the slope of the methane mixing ratio against time and to throw away data that yield a linear regression value of $r^2 < 0.95$. On the basis of the regression an uncertainty due to measurement errors is estimated to be <5%, but it can be questioned whether the closed chamber approach does not induce intrinsic errors. So far, the debate on that issue is inconclusive. An important condition for a small error is that seasonal field measurements are used. *Koyama* [1963] estimated methane emissions from 1 month laboratory incubation experiments and obtained highly different estimates than others.

[13] A second source of uncertainty is the lack of quantitative understanding of processes underlying methane emissions. Examples are the limited knowledge on root gas transport capacity [*Segers and Rappoldt*, 2000], interactions between methanogens and iron reducers [*Ratering and Conrad*, 1998], and the microbial dynamics of methanotrophs [*Roslev and King*, 1995]. The influence of, e.g., different descriptions of methane oxidation dynamics on methane emission estimates is considerable [*van Bodegom et al.*, 2001a]. It is, however, not possible to quantify the uncertainty introduced by this lack of knowledge unless new experiments are performed on those processes.

[14] A third source of uncertainty is the simplification of processes in the model. This uncertainty can be quantified by comparison with a fully mechanistic model. *Segers et al.* [2001] showed that the influence of neglecting spatial heterogeneities in microbial kinetics was small. Also, the simplifications made in the model used in this case study can be compared with results obtained from fully mechanistic models. A comparison of the description for methane production with a mechanistic methane production model [*van Bodegom and Scholten*, 2001] showed that the seasonal methane production rates were not significantly different ($P = 0.85$) for the two models at the validation conditions tested by *van Bodegom et al.* [2001a]. Indeed, the dynamics during the season were different for the two models (data not shown), but that does not influence the upscaling estimates since we use seasonally integrated emission estimates.

[15] A final source of uncertainty is the estimation of input parameter values. Apart from the question whether an input parameter value is representative for a site (which is treated in section 3), even if a limited number of input variables is required, these cannot be derived directly from databases. A highly relevant example is the limited, if not absent, information on the use of organic fertilizer input, which largely influences emissions [*Denier van der Gon and Neue*, 1995]. Another example is the estimation of the amount of reducible iron. In such occasions, these input parameter values have to be deduced from other attributes. This deduction of a property adds to the uncertainty in the final result. Reducible iron, for example, needs to be determined from data on extracted iron, present in, e.g., chemical databases belonging to soil maps. The effect of these so-called data transfer functions depends on the sensitivity of the model to these data. This model sensitivity is, however, not constant and depends on the values of other input parameters but is generally high for the two examples given [*van Bodegom et al.*, 2000]. We will return to the importance of data in section 3.1.

[16] *van Bodegom et al.* [2001a] investigated the deviation between modeled and measured seasonal methane emissions, which was surprisingly small and insignificant given all uncertainties. The coefficient of variation introduced by modeling the methane emissions was 7% of the measured seasonal emission rate. The differences between diel modeled and measured methane emissions were larger. Their model could not explain this small temporal scale variability, but that does not affect the seasonally integrated methane emission prediction. We can thus conclude that

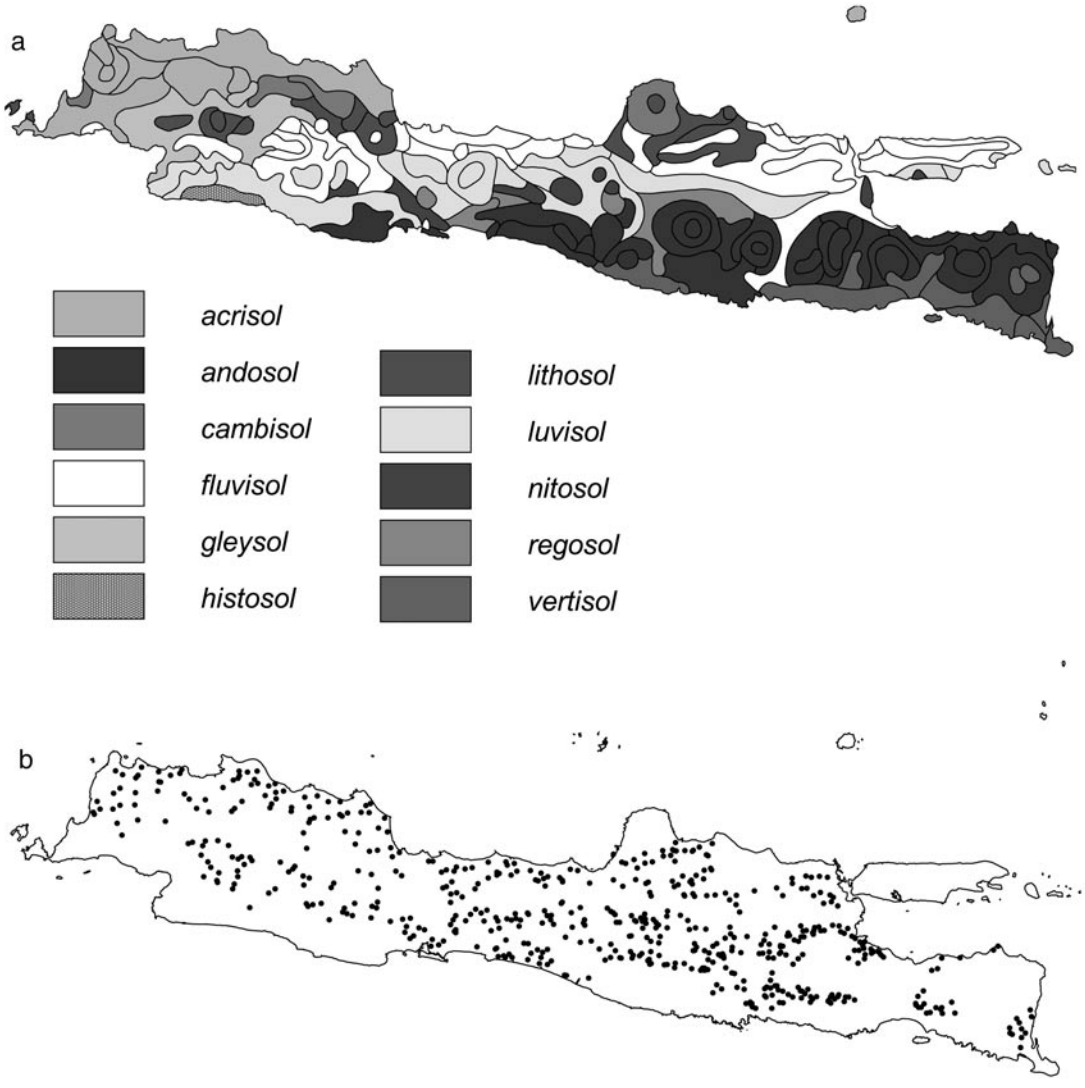

Figure 2. The data sources: (a) The FAO 1:5,000,000 soil map for Java and (b) the spatial distribution of soil samples included in the CSAR data set.

for the model used in this paper the uncertainties induced by modeling are small at the plot scale.

3. Data and Upscaling

3.1. Data Uncertainties

[17] Several uncertainties surround data, one of which (the uncertainty induced by data transfer functions) was briefly introduced in section 2.3. The first uncertainty concerns the data availability. Spatial information on soil properties is usually derived from soil maps. The Food and Agriculture Organization (FAO) 1:5,000,000 soil map of the world is used most frequently in upscaling studies on methane [see, e.g., *Bachelet and Neue*, 1993; *Cao et al.*, 1998]. However, classification of soil maps in rice-growing regions, the FAO soil map included, is often heavily based on geomorphological characteristics rather than on soil sampling and successive laboratory analysis. Not surprisingly, soil organic matter and alternative electron acceptors under anaerobic conditions, which are the key variables in the driving processes for methane emissions, are not discriminative elements between soil types or mapping units. These soil chemical variables have to be extracted from soil type information, with the potential risk that the variability of these variables within a soil type is considerable and may be larger than between soil types. A second potential risk with soil maps as a source for soil chemical data is that soil types are assumed to be independent of current land use, whereas topsoil properties do change under influence of land use, especially by rice cultivation. Average soil chemical characteristics for a region may thus not represent soil chemical characteristics of rice paddies within that region.

[18] For the case study of Java we had access to two different soil data sources: the FAO 1:5,000,000 soil map for Java (Figure 2a) and soil samples of Java which had been collected and analyzed on soil chemical and morphological properties by the Center for Soil and Agroclimate Research (CSAR) in Bogor, Indonesia. The 555 soil samples in the CSAR data set (Figure 2b) were taken from rice fields throughout Java and included all important rice growing areas on Java, with the exception of the island of Madura at the eastern side of Java. Thus the CSAR data did not have the disadvantages associated with the FAO soil map and similar soil maps for deriving rice soil chemical properties. On the other hand,

the CSAR data set is only available for Java, and such detailed data sets are virtually nonexistent, while methane emission estimates are most relevant at the national and global scale. In this case study, estimates from both data sources are described and compared, and the effect of data source selection is quantified.

[19] Deriving soil chemical properties from the FAO soil map is a difficult, assumption-laden process: The polygons of the FAO soil map are mapping units and not unique soil types. Each mapping unit consists of one or two dominant soil types and one or more associated soil types. However, the location of the soil types within the mapping unit is unknown, complicating the linkage to rice areas. This implies that average soil chemical properties for the mapping units have to be derived from the soil types using general soil type profile data. *Denier van der Gon et al.* [2000] and *Knox et al.* [2000] describe this procedure in more detail. Soil organic carbon for the various soil types was estimated from soil profile data in the World Inventory of Soil Emission potentials (WISE) database [*Batjes*, 1995]. Soil organic carbon was directly available for all samples in the CSAR data set. Soil samples with soil organic carbon >15% and histosols were excluded from further analysis because the methane emission model was developed for mineral soils. Reducible iron had to be estimated from iron extraction data, which introduces a third data uncertainty, the application of data transfer functions. Iron extraction data belonging to the FAO soil map are available in the ISIS database *van de Ven and Tempel*, 1994]. We used information on oxalate extractable iron, which is a measure of amorphous iron oxides, that can be reduced well [*Lovley*, 1991] and dithionite extractable iron. Dithionite extractable iron combines amorphous and crystalline iron oxides. Only a few percent of the crystalline iron oxides can be reduced [*Lovley and Phillips*, 1986; *Roden and Zachara*, 1996]. From the combination of these two iron extractions the total amount of reducible iron can be estimated accurately because other iron minerals are even less reactive [*Raiswell et al.*, 1994]. The CSAR database only contained information on dithionite extractable iron. Dithionite extractable iron was directly related to reducible iron using the relationship given by *Roden and Zachara* [1996].

3.2. Data Scale

[20] It is not possible to collect the input variables at the scale at which the emission model was developed, i.e., for a single rice plant. To approach this scale as closely as possible, it is necessary to choose the highest data resolution possible for the input data. In other words, data availability determines the resolution for upscaling, and well-documented regions are thus of prime importance. The model needs data on soils and agricultural and climatic aspects. The scale at which soil information is available is described above. For Java, agricultural data are available at a district scale (called kabupaten/kotamadya), and climatic data were available at a similar scale [*Verburg et al.*, 1999]. As a result, the highest data resolution possible was 20 × 20 km, leading to 329 grid cells to describe Java. This has major implications for the upscaling results, as will be shown in section 4.3.

3.3. Data Interpolation

[21] All data must become available at the data scale determined above. For point data (like in the CSAR database, Figure 2b) these data have to be interpolated to match the grid size of the data scale. In order to present the spatial heterogeneity correctly, interpolation techniques use the autocorrelation of variability by the application of semivariograms. Different models (spherical, exponential, linear with sill, or Gaussian models, etc.) can be applied to quantify the semivariogram. In the case study the best model for organic carbon was a spherical model, and for dithionite extractable iron the best model was an exponential model [*van Bodegom et al.*, 2001b].

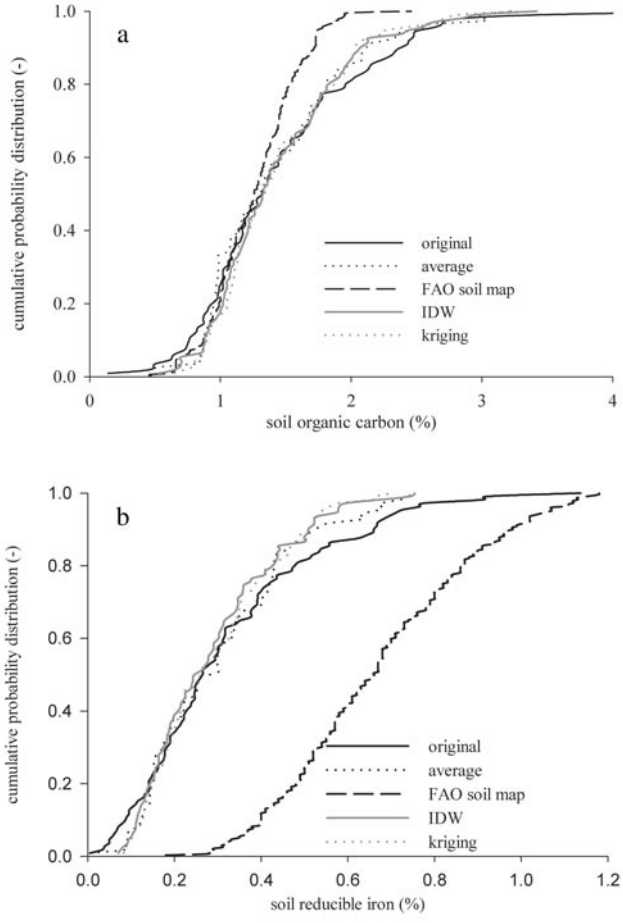

Figure 3. Cumulative probability distribution functions (pdf) for (a) organic carbon and (b) soil iron content under influence of different decisions made for interpolation of soil properties.

[22] Given a semivariogram model and a set of parameters, interpolation to unsampled locations can be obtained through kriging. The number of neighbors was optimized by a cross-validation procedure [*Voltz and Webster*, 1990], and the number was different for organic carbon and dithionite extractable iron [*van Bodegom et al.*, 2001b]. The interpolation results obtained from optimized ordinary kriging were, although statistically better [*van Bodegom et al.*, 2001b], hardly different from results using inverse distance weighing (IDW) interpolation with an optimized number of neighbors [*van Bodegom et al.*, 2001b] or using average grid values instead of interpolation (Figure 3). By using interpolation some information on spatial variability is lost, as is indicated by the steeper probability density functions for interpolation compared to the original data (Figure 3). A reliable probability function for input parameters is important because errors might be introduced by the combination of interpolation with a nonlinear model response to input parameters (section 4.2).

[23] In a case where a soil map is used as a data source (in our case the FAO 1:5,000,000 soil map) the mapping units of the soil map have to be averaged (weighed for the area occupied by a mapping unit) to the grid size of the target scale. Averaging only makes sense if the variability between soil units is larger than within soil units and if the characteristics of a soil unit are statistically different from each other, as was the case in our study

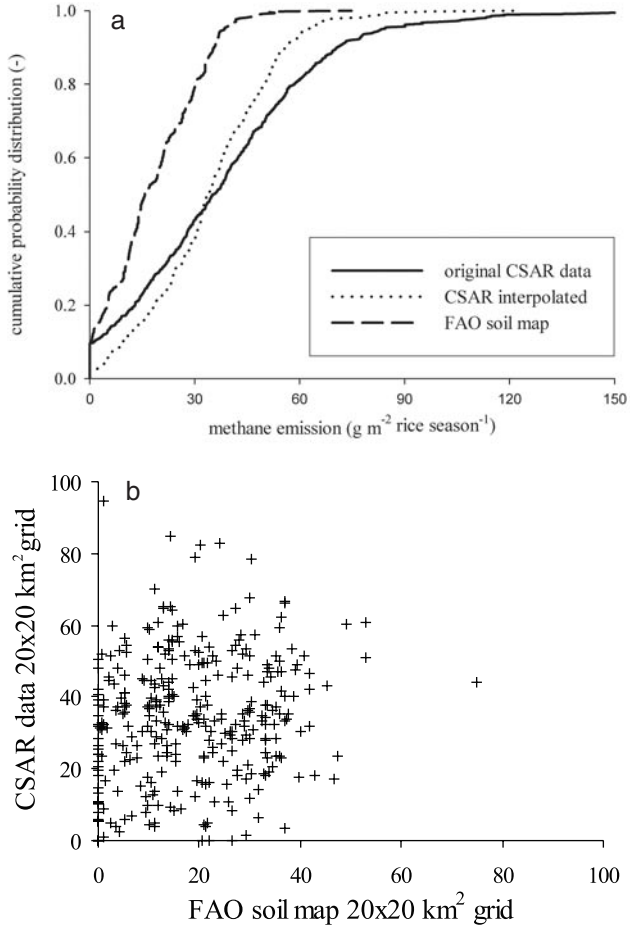

Figure 4. Effects of data source and interpolation on seasonal methane emission estimates (g m^{-2} $season^{-1}$) for (a) probability distribution function for methane for kriging with CSAR interpolated soil map and the original CSAR data and (b) correlation between emission estimates for 20 × 20 km grids for both data sources.

(results not shown). Since we were only interested in rice paddy soils, soils for which it seemed highly probable that there was no wetland rice production possible were not included in the analysis. The selection of soil units that are presumably used for growing rice on Java is based on *Soepraptohardjo and Suhardjo* [1978]. An implicit assumption of the use of reference profiles as obtained from *Batjes* [1995] is that soil chemical properties have not been significantly modified by rice production, although this does not seem probable [*Denier van der Gon et al.*, 2000]. However, there is no alternative because the available number of rice soil profiles is far too small to cover all relevant soil types.

[24] Methane emission estimates were calculated varying only the soil data input, keeping all other input parameters at their average values given in Table 1. The deviation in methane emission estimates due to different semivariogram models or a different number of neighbors in the interpolation of soil point data to the grid size of 20 × 20 km was <1% [*van Bodegom et al.*, 2001b]. Information on the spatial heterogeneity of soil properties and of estimated methane emissions was, however, lost by the interpolation of soil properties. The variance in methane emission estimates proved to be significantly different ($P < 0.001$) from the variance expressed using the original CSAR database for the soil properties [*van Bodegom et al.*, 2001b] (Figure 4a).

[25] The effects of the data source on emission estimates at a grid size of 20 × 20 km are considerably larger (Figure 4). Both the average and the variance in calculated methane emissions are significantly different (at $P < 0.05$), as shown by a Student's t test and an F test. This is caused by the significant differences (at $P < 0.05$) in the interpolated soil properties from both data sources. The differences between the data of the two data sources are so large that there is no clear relation between the emission estimates for the individual grids (Figure 4b).

4. Scales and Upscaling

4.1. Lateral Interactions

[26] The occurrence of lateral interactions needs to be considered before the model and the (interpolated) data can be combined to estimate regional methane emissions. Lateral interactions are spatial interactions between grids, implying that the characteristics of the output variable(s) are scale dependent. Such interactions occur, for instance, in various hydrological studies [e.g., *Dunne et al.*, 1991]. For greenhouse gas emissions it is usually assumed that such lateral interactions are not important [*King et al.*, 1989].

[27] This assumption was tested for methane emission estimates. As indicated above, methane emission from rice paddies is determined by methane production, oxidation, and transport. Methane production is a microbial process that occurs locally within the soil and is thus spatially independent. Methane transport in rice paddies is diffusion driven and is, in principle, spatially dependent because it depends on local atmospheric methane concentrations. Methane concentration differences between soil and atmosphere are, however, very large, and this scale effect can be neglected. Soil methane oxidation cannot occur at atmospheric methane concentrations [*Hanson and Hanson*, 1996], and lateral interactions from this process are not expected. Methane oxidation can, however, also occur on or within aboveground rice biomass using atmospheric methane. This methane oxidation is a diffusion-limited process and thus, in principle, laterally affected by local atmospheric methane concentrations.

[28] Methane consumption kinetics in this environment were quantified by sampling fresh rice leaves, cutting these into pieces of 1 cm, and incubating them in a sterile nitrate mineral salts (NMS) medium [*Whittenbury et al.*, 1970]. At 10% methane in the headspace, methane was consumed (Figure 5a), proving that methane oxidation is potentially possible in this environment. However, at a 10 ppm methane concentration, as is normally encountered within a canopy, no significant methane oxidation could be detected (Figure 5b). Therefore lateral interactions can be neglected.

4.2. Nonlinear Model Responses

[29] A process-based model, as the one presented above, combines several nonlinear processes and interactions. It might therefore be expected that the model output also reacts nonlinearly to input parameters. In the case of a nonlinear model response the use of arithmetic means of an input parameter does not lead to a true average output value. This is often referred to as the "fallacy of averages" [*Rastetter et al.*, 1992]. In the present study we focused on effects of spatially explicit soil information, in particular, the effects of reducible iron and organic carbon (see section 3). The correlation between these input parameters and modeled methane emissions was determined for the two data sources at a data resolution of 20 × 20 km. Figure 6 shows that the relationship between the soil input parameters and modeled methane emissions is, although significant for both data sources ($P < 0.05$), highly

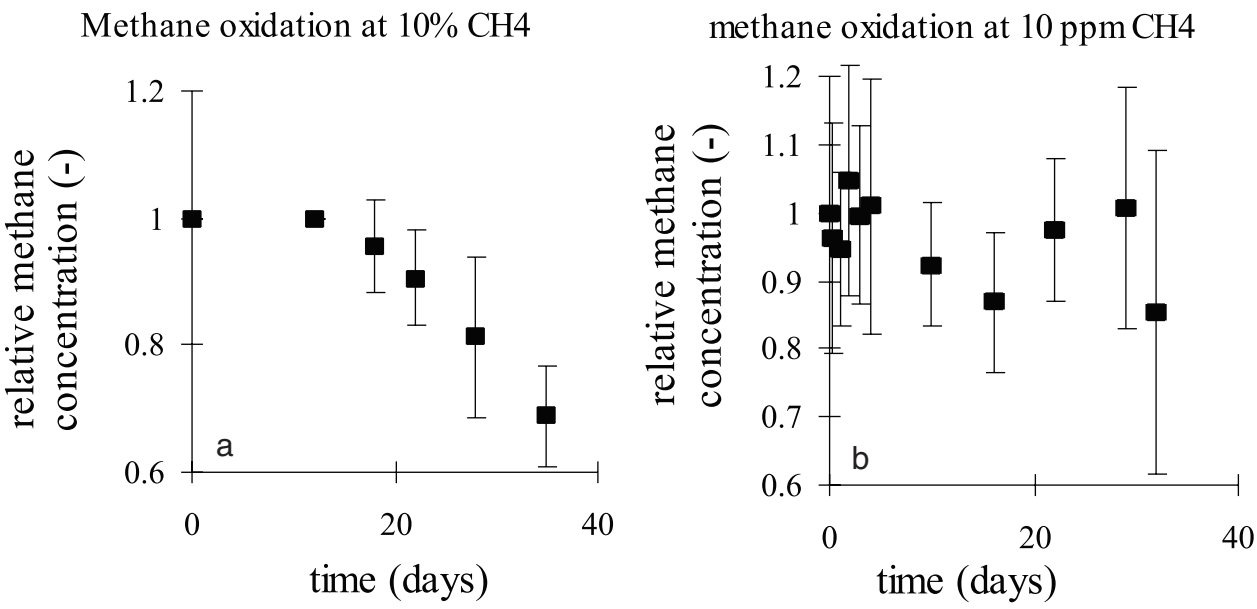

Figure 5. Methane oxidation at aboveground rice biomass measured in incubation experiments with (a) 10% methane and (b) 10 ppm methane in the headspace.

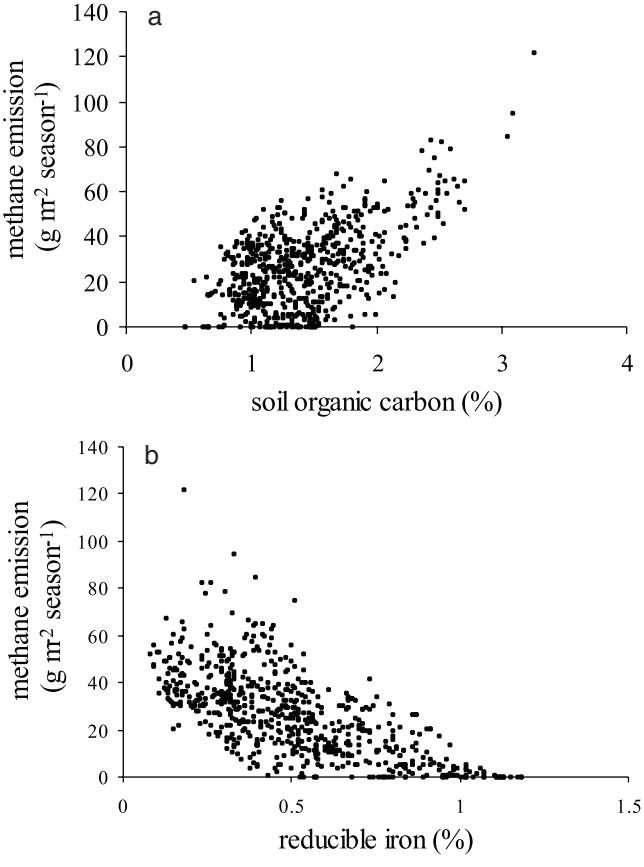

Figure 6. Model estimation of methane emissions under influence of the FAO 1:5,000,000 soil map soil organic carbon and the CSAR data set at Java for (a) soil organic carbon and (b) soil reducible iron content.

scattered and nonlinear, especially for soil reducible iron. The relationship with soil organic carbon is more linear, although the underlying processes would suggest a nonlinear relationship. Part of the nonlinearities might have been compensated by the highly significant ($P < 0.001$) correlation between soil organic carbon and soil reducible iron in both data sources.

4.3. Data Resolution Effects

[30] The minimal grid size of 20 × 20 km for the upscaling of methane emission estimates in the case study of Java was determined by the data availability (section 3.2), while processes leading to methane emissions occur at a more detailed scale (<1 m^2). Given the nonlinear response of the model (section 4.2), lumping errors due to the use of data with a resolution of 20 × 20 km can be significant [*King*, 1991]. Many upscaling studies on methane emission estimates do, however, use even lower resolutions of 1° × 1°, ~100 × 100 km [*Bachelet and Neue*, 1993], or 0.5° × 0.5°, ~50 × 50 km [*Cao et al.*, 1998]. Unfortunately, the question about the optimal data resolution for the calculating methane emissions has never been answered properly.

[31] To test such effects of data resolution, the kriged data from the CSAR data set and from the FAO soil map were gridded additionally to 40 × 40 km and 100 × 100 km (~1° × 1°) using the original data set. A different outcome is expected when calculating emissions first, followed by averaging the results (i.e., at 20 × 20 km), than when averaging the data first, followed by calculating emissions (i.e., at 100 × 100 km). Given the fact that there are no lateral interactions, the model was run for all grids independently at all three data resolutions for both soil data sources. The grid estimates were corrected for the percentage under rice cover.

[32] The grid averages of reducible iron and soil organic carbon did not change significantly (at $P < 0.05$) by kriging at different data resolutions either for the CSAR data or for the FAO data. The grid mean variance [*van Bodegom et al.*, 2001b], however, decreased significantly for the different data resolutions for both the CSAR data set and the FAO data set, as shown by an F test at $P < 0.05$. This decreased variance in combination with the nonlinear model responses led to both a significant change in the average calculated methane emission ($P < 0.05$) and in a

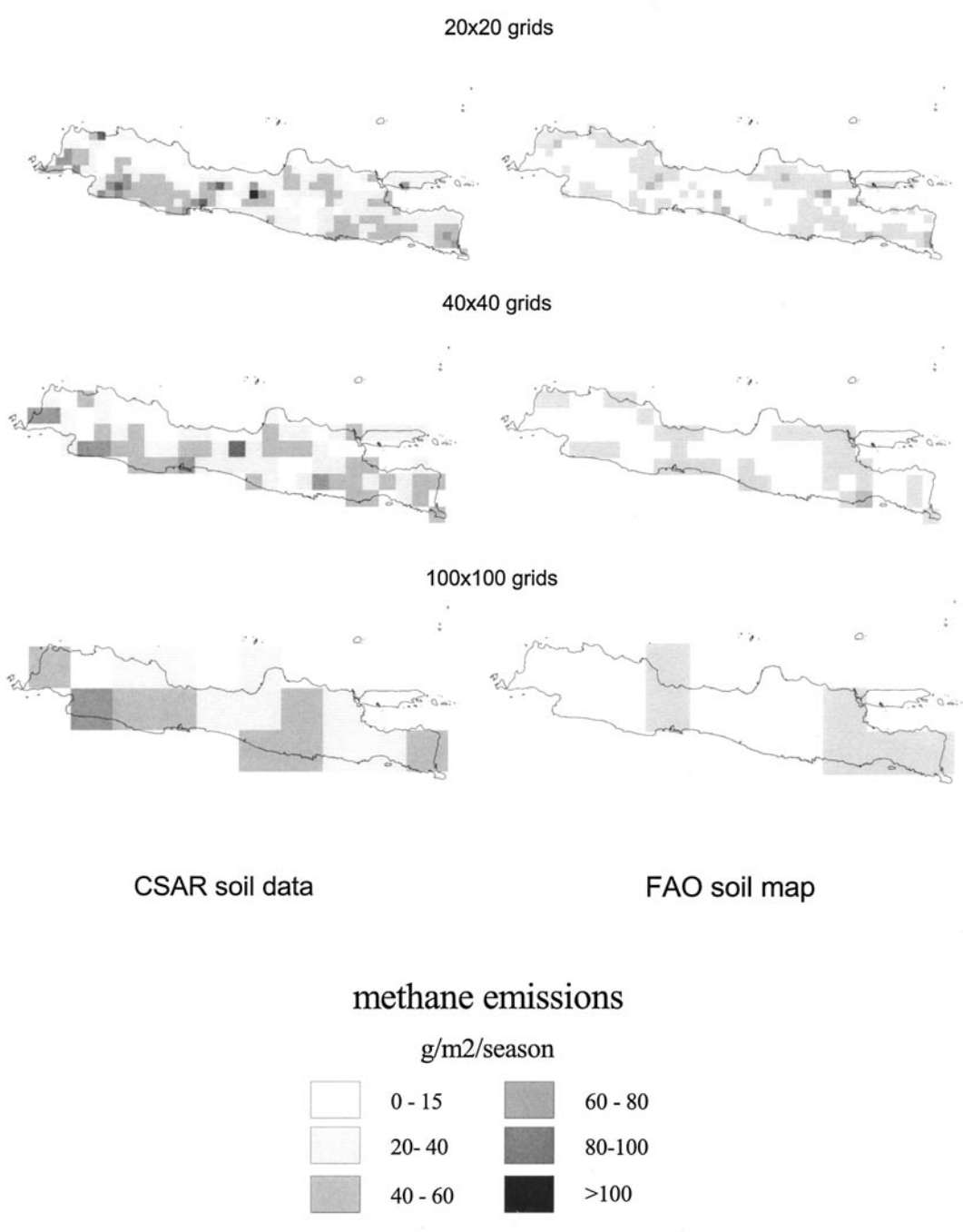

Figure 7. Effects of data resolution on the estimate of methane emissions, expressed in g CH_4 m^{-2} rice $season^{-1}$, for Java with 20 × 20 km grids, 40 × 40 km grids, and 100 × 100 km grids for the CSAR data set and the FAO soil map, respectively.

significant decrease in the variance in calculated methane emission ($P < 0.05$) for both data sources, as shown by Student t tests and F tests. Not only did the average calculated methane emission change under influence of data resolution effects, but also the methane emission estimate from individual grids was significantly different (at $P < 0.05$) at coarser data resolutions compared to the average emission estimate for that grid at the 20 × 20 km resolution. Methane emission estimates at different data resolutions are shown in Figure 7.

[33] The results show that (1) the average emission estimate changes going from detailed to coarser data resolutions, (2) the spatial variability in methane emission estimates decreases, especially at a coarser data resolution, and (3) different data sources lead to different data resolution effects under influence of the interaction between data source and model. These scaling effects only occur because the model responds in a nonlinear fashion to data. The results of this specific situation can thus not be extrapolated to other systems with a different data set and/or a different

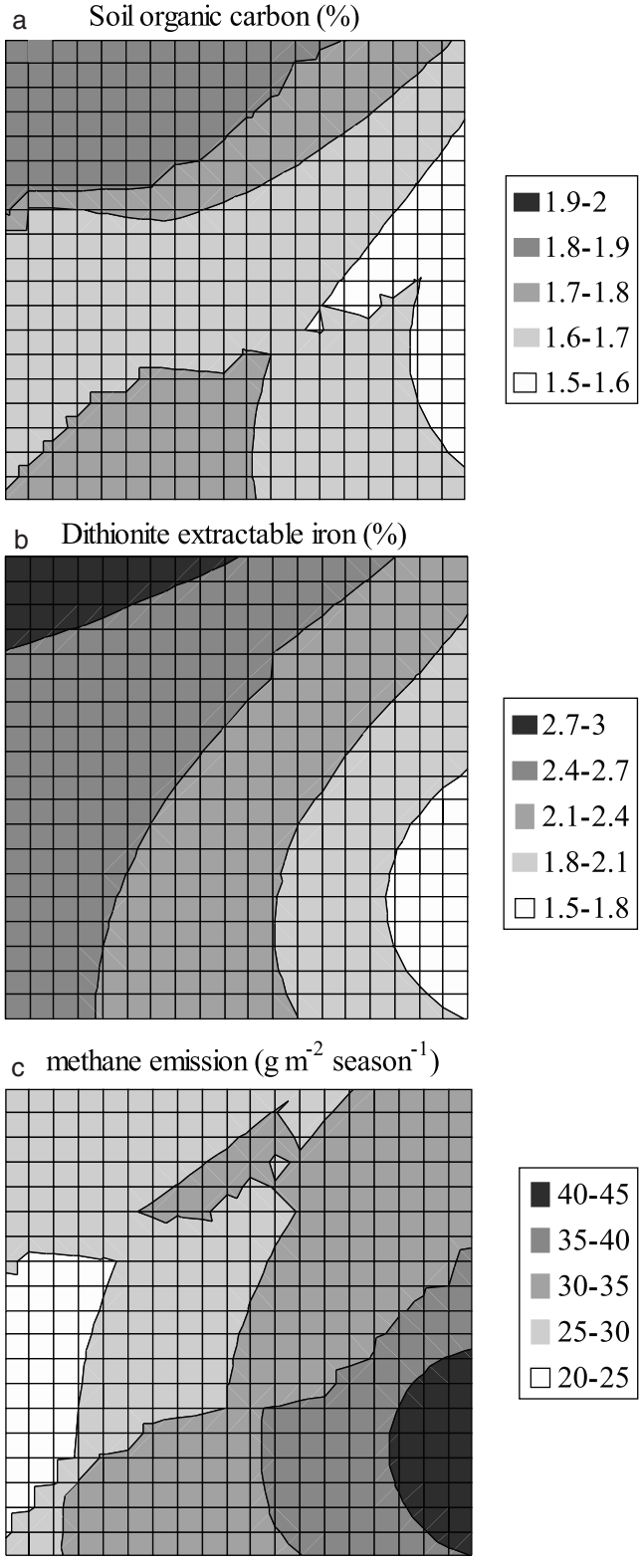

Figure 8. Spatial variability in (a) kriged soil organic carbon (b) kriged dithionite extractable iron, and (c) calculated methane emissions for one grid cell of 20 × 20 km.

model. The strong loss of information on spatial variability in calculated methane emission patterns with coarser resolution (stronger than the effects on average calculated methane emissions) suggests that data resolution for this type of study should always be kept as high as possible.

[34] In section 3.3 it was shown that substantial data variance was lost owing to interpolation. In combination with the nonlinear model response (section 4.2) and the gap between the process-based field-scale emission model and the resolution determined by data availability, one would expect a bias in methane emission estimates already at the 20 × 20 km grid resolution.

[35] To test the effects of higher resolutions, one grid cell was chosen randomly as a case study and subdivided in 1 × 1 km grids. The kriged values for dithionite extractable iron and soil organic carbon from the CSAR data set at the data resolution of 1 × 1 km were used as input for the model, and methane emissions were calculated for these individual grid cells. Not only were methane emission estimates highly variable even within this grid cell (Figure 8), but the average methane emission calculated from these 400 grids was also different by 10% from the methane emission calculated from the 20 × 20 km data resolution. The 20 × 20 km data resolution indeed seems to bias the estimates of methane emissions already. Unfortunately, it will not be possible to calculate methane emissions at the higher data resolution on a regular basis because other input variables are not known at such a high data resolution. However, part of this resolution problem might be solved by the application of Monte Carlo techniques to calculate the variability in the estimate.

5. Discussion

[36] Different factors that influence the outcome of methane emission estimates by upscaling were evaluated with a case study on Java. Even though we focused on the effects by soil properties in this case study, keeping other input parameters at a constant average value, the scaling effects were considerable. We distinguished effects under influence of model, data, and scale. The interactions between data and scale are summarized in Figure 1, and on top of the resulting individual methane emissions estimates, effects by lateral interactions and data resolution are possible. All effects on average methane emission estimates are summarized in Table 2. In addition, a relative change in estimated methane emission under influence of one of the factors in comparison to a situation of 20 × 20 km grids and optimized ordinary kriging based on the CSAR data set is shown.

[37] The effects on the estimates by different factors differ considerably in this case study. The influence of data is the largest, followed by effects of scale (Table 2). These effects were all significant (at $P < 0.05$).

[38] The uncertainties induced by the model were the smallest. These small uncertainties might partly be caused by the fact that only seasonal emissions were considered and that we made no comparison with other field-scale models. The comparison with field emission data and with a fully mechanistic model does, however, give some fidelity in the accuracy of a model. In principle, it must be possible to apply a process-based model developed for the field scale at the district scale because no new processes occur at a larger scale. It remains, however, problematic that results obtained at a district scale cannot be validated by direct measurements.

[39] The effects of the selection, choice, derivation, and use of data are much larger. This shows that it is difficult to provide the model with accurate input data. Especially, the effects of soil data source are large. This might be due partly to the use of different soil analysis methodologies, partly to the use of a soil map instead of direct chemical analyses, and partly to the fact that data had to be summarized/interpolated to a different scale than they originated.

Table 2. Scaling Effects on Average Emission Estimate and Included Spatial Heterogeneity (Calculated From the Coefficient of Variation) as a Result of Various Influences on CH_4 Emission Per Surface Area Planted With Rice at Java

	Average Emission, g m^{-2} rice $season^{-1}$	Spatial Heterogeneity, %	Difference From Default, %
Default estimate	34.5	53	0
Effects of model			
Model uncertainties	n.d.	n.d.	7[a]
Full versus simplified model	n.d.	n.d.	12[b]
Effects of data			
FAO soil map versus CSAR set	18.1	72	48
No interpolation versus kriging	23.6	63	32
IDW versus kriging	34.5	51	<1
Effects of scale			
Lateral interaction	n.d.	n.d.	n.d.
1 × 1 resolution	n.d.	15.5[c]	10.0[c]
40 × 40 resolution[d]	34.9, 16.9	51, 62	1, 51
100 × 100 resolution[d]	37.5, 16.4	43, 47	9, 52
Resolution of whole Java[d]	100.4, 15.6	0, 0	191, 55

[a] Calculated for nine sites throughout Southeast Asia, one of which is on Java.
[b] Calculated for the methane production approximation for the sites mentioned in footnote a.
[c] Calculated only for the one grid cell under study.
[d] The first of the pair of values is for CSAR data set, and the second is for FAO.

[40] These scaling effects are prominent, even though a very high default resolution of 20 × 20 km was chosen, and account for an uncertainty of 10–20% (based on the CSAR data set). It seems thus of importance to consider these scaling effects for global methane emission estimates when resolutions of 0.5° × 0.5° or 1° × 1° are used. Even a 20 × 20 km resolution might not be detailed enough to obtain unbiased emission estimates, and the scaling effect might still be underestimated. A resolution of 1 × 1 km gave a 10% different emission estimate, and it seems possible that even higher biases might occur because the within-field variability in methane emissions typically ranges from 7 to 30% [*Denier van der Gon and Neue*, 1995] with extremes up to 80% [*Wassmann et al.*, 1996]. Field to field variation might lead to additional variability in methane emission from rice paddies [*Khalil et al.*, 1998]. Moreover, in this case study, only the effect due to soil properties was determined, and inclusion of other input parameters might further increase variability and scaling effects. The number of measurements/estimates is thus more important than the accuracy of one measurement to estimate regional methane emissions. The variability that causes the effects of data interpolation and data resolution might be captured only with many predictions (even if there is no spatial dependency between the data).

[41] Many of the scaling effects could only occur because of model, data, and scale interactions. Therefore average input values do not correspond necessarily to average output values. Information on variability was lost during upscaling because models decrease the variability as models describe general trends only, and this variability was decreased further under influence of data interpolation and data resolution. Scaling effects are therefore situation dependent and influenced by the combination of model, data handling, scale chosen, and site variability. Given the scaling effects, it might be questioned whether an unbiased methane emission estimate is possible, given the principal data limitations.

[42] Under influence of upscaling, the effects on average methane emission estimates are different from the effects on total emission estimate, even though the total amount of rice area was kept constant for all situations (compare Tables 2 and 3). This is probably again due to changes in spatial heterogeneity, through which high methane emissions are allocated to different sites in different situations. Note that the total emission estimate does not represent the actual methane emission from rice fields on Java. Other estimates may arise if also the spatial variability in other driving variables is considered.

6. Conclusions

[43] Clear scaling effects were observed in the presented case study for estimates on methane emissions from rice paddies. Model uncertainties and data interpolation effects introduced small limitations to the emission estimate, and data accuracy was limiting the

Table 3. Influence of Scaling Effects as a Result of the Various Factors on Total CH_4 Emission Estimate for Java

	Total Emission Estimate, Tg	Difference From Default, %
Default estimate	1.67	n.d.
Effects of data		
FAO soil map versus CSAR set	0.95	43
No interpolation versus kriging	1.97	18
IDW versus kriging	1.68	<1
Effects of scale		
40 × 40 resolution[a]	1.72, 0.89	3, 47
100 × 100 resolution[a]	2.00, 0.86	19, 49
Resolution of whole Java[a]	5.15, 0.80	208, 52

[a] The first of the pair of values is for CSAR data set, and the second is for FAO.

emission accuracy. Many of the scaling effects described in this study are not often accounted for in upscaling studies for the estimation of global methane emissions and might have biased the present estimates (which might also present part of the explanation for the large uncertainty in the global estimates). As the processes leading to methane emissions occur at the scale of a single plant, more detailed databases of relevant and sensitive input parameters are needed to summarize the outcome of the processes more properly and to avoid scaling effects in global upscaling studies. Especially, more detailed data on organic fertilizer management, water management, and soil organic carbon are needed, although it will never be possible to collect data at the scale of the processes (which would be ideal to avoid scaling effects). However, this type of information is not widely collected and is hard to obtain. As long as such information is not available, it will be hard to obtain emission estimates that have a lower uncertainty than the presently generated results. A dialogue between modelers of greenhouse gas emission processes and ''data collectors'' should be stimulated to provide the latter with motivation and insights that particular additional data, which sometimes might be collected for small additional costs, may greatly enhance the value of existing data sets for other purposes.

[44] **Acknowledgments.** We thank the Interregional Research Programme on Methane Emissions from Rice Fields' and especially R. Wassmann, R. S. Lantin, A. K. Makarim, and P. Setyanto for cooperation, discussion, and sharing of unpublished results. N. H. Batjes (ISRIC) is gratefully thanked for making the WISE database available and supplying a subset of the ISIS database containing soil Fe data. The CSAR data set was kindly provided by Sri Adiningsih. The research was supported financially by the Dutch National Research Program on Global Air Pollution and Climate Change.

References

Bachelet, D., and H. U. Neue, Methane emissions from wetland rice areas of Asia, *Chemosphere*, *26*, 219–237, 1993.

Batjes, N. H., (Ed.), A homogenized soil data file for global environmental research: A subset of FAO, ISRIC and NRCS profiles (version 1.0), *Working Pap. 95/10*, Int. Soil Ref. and Inf. Cent. (ISRIC), Wageningen, Netherlands, 1995.

Biro Pusat Statistik (BPS), Struktur ongkos usaha tani padi dan palawija, Jakarta, Indonesia, 1994.

Biro Pusat Statistik (BPS), Survei partanian-Produksi tanaman padi di Indonesia 1994, Jakarta, Indonesia, 1995.

Biro Pusat Statistik (BPS), Survei luas tanaman padi dengan pendekatan rumah tangga di Jawa 1996, Jakarta, Indonesia, 1996.

Cao, M., K. Gregson, and S. Marshall, Global methane emission from wetlands and its sensitivity to climate change, *Atmos. Environ.*, *32*, 3293–3299, 1998.

Choularton, T. W., M. W. Gallagher, K. N. Bower, D. Fowler, M. Zahniser, and A. Kaye, Trace gas flux measurements at the landscape scale using boundary-layer budgets, *Philos. Trans. R. Soc. London*, *351*, 357–368, 1995.

Conrad, R., Mechanisms controlling methane emission from wetland rice fields, in *Biogeochemistry of Global Change: Radiatively Active Trace Gases*, edited by Oremland, R. S., pp. 317–335, Chapman and Hall, New York, 1993.

Denier van der Gon, H. A. C., and H. U. Neue, Influence of organic matter incorporation on the methane emission from a wetland rice field, *Global Biogeochem. Cycles*, *9*, 11–22, 1995.

Denier van der Gon, H. A. C., P. M. van Bodegom, S. Houweling, P. H. Verburg, and N. van Breemen, Combining upscaling and downscaling of methane emissions from rice fields: Methodologies and preliminary results, *Nutr. Cycles Agroecosyst.*, *58*, 285–301, 2000.

Dunne, T., W. H. Zhang, and B. F. Aubry, Effects of rainfall, vegetation, and microtopography on infiltration and runoff, *Water Resour. Res.*, *27*, 2271–2285, 1991.

Hanson, R. S., and T. E. Hanson, Methanotrophic bacteria, *Microbiol. Rev.*, *60*, 439– 471, 1996.

Holzapfel-Pschorn, A., and W. Seiler, Methane emission during a cultivation period from an Italian rice paddy, *J. Geophys. Res.*, *91*, 11,803–11,814, 1986.

Houghton, J. T., L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Marskell, *Climate change 1995: The Science of Climate Change*, Cambridge Univ. Press, New York, 1996.

Khalil, M. A. K., R. A. Rasmussen, M. J. Shearer, R. W. Dalluge, L. X. Ren, and C.-L. Duan, Measurements of methane emissions from rice fields in China, *J. Geophys. Res.*, *103*, 25,181–25,210, 1998.

King, A. W., Translating models across scales in the landscape, in *Quantitative Methods in Landscape Ecology*, edited by Turner, M. G. and R. H. Gardner, pp. 479–517, Springer-Verlag, New York, 1991.

King, A. W., R. V. O'Neill, and D. L. Deangelis, Using ecosystem models to predict regional CO_2 exchange between the atmosphere and the terrestrial biosphere, *Global Biogeochem. Cycles*, *3*, 337–361, 1989.

Knox, J., R. Matthews, and R. Wassmann, Using a crop/soil simulation model and GIS techniques to assess methane emissions from rice fields in Asia, II, Databases, *Nutr. Cycles Agroecosyst.*, *58*, 179–199, 2000.

Koyama, T., Gaseous metabolism in lake sediments and paddy soils and the production of atmospheric methane and hydrogen, *J. Geophys. Res.*, *68*, 3971–3973, 1963.

Lelieveld, J., P. J. Crutzen, and F. J. Dentener, Changing concentration, lifetime and climate forcing of atmospheric methane, *Tellus, Ser. B.*, *50*, 128–150, 1998.

Lovley, D. R., Dissimilatory Fe(III) and Mn(IV) reduction, *Microbiol. Rev.*, *55*, 259–287, 1991.

Lovley, D. R., and E. J. P. Phillips, Organic matter mineralization with reduction of ferric iron in anaerobic sediments, *Appl. Environ. Microbiol.*, *51*, 683–689, 1986.

Neue, H. U., Fluxes of methane from rice fields and potential for mitigation, *Soil Use Manage.*, *13*, 258– 267, 1997.

Nouchi, I., T. Hosono, K. Aoki, and K. Minami, Seasonal variation in methane flux from rice paddies associated with methane concentration in soil water, rice biomass and temperature, and its modeling, *Plant Soil*, *161*, 195– 208, 1994.

Raiswell, R., D. E. Canfield, and R. A. Berner, A comparison of iron extraction methods for the determination of degree of pyritisation and the recognition of iron-limited pyrite formation, *Chem. Geol.*, *111*, 101–110, 1994.

Rastetter, E. R., A. W. King, B. J. Cosby, G. M. Hornberger, R. V. O'Neill, and J. E. Hobbie, Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems, *Ecol. Appl.*, *21*, 55–70, 1992.

Ratering, S., and R. Conrad, Effects of short-term drainage and aeration on the production of methane in submerged rice soil, *Global Change Biol.*, *4*, 397–407, 1998.

Roden, E. E., and J. M. Zachara, Microbial reduction of crystalline iron (III) oxides: Influence of oxide surface area and potential for cell growth, *Environ. Sci. Technol.*, *30*, 1618– 1628, 1996.

Roslev, E., and G. M. King, Aerobic and anaerobic starvation metabolism in methanotrophic bacteria, *Appl. Environ. Microbiol.*, *61*, 1563–1570, 1995.

Rosswall, T., R. G. Woodmansee, and P. G. Risser (Eds.), *Scale and Global Changes: Spatial and Temporal Variability in Biospheric and Geospheric Process*, *SCOPE 35*, John Wiley, New York, 1988.

Segers, R., Methane production and methane consumption: A review of processes underlying wetland methane fluxes, *Biogeochemistry*, *41*, 23–51, 1998.

Segers, R., C. Rappoldt, and P. A. Leffelaar, Modeling methane fluxes in wetlands with gas transporting plants, 2, Soil layer scale, *J. Geophys. Res*, *106*, 3541–3558, 2001.

Simpson, I. J., G. W. Thurtell, G. E. Kidd, M. Lin, T. H. Demetriades-Shah, I. D. Flitcroft, E. T. Kanemasu, D. Nie, K. F. Bronson, and H. U. Neue, Tunable diode laser measurements of methane fluxes from an irrigated rice paddy field in the Phillipines, *J. Geophys. Res.*, *100*, 7283–7290, 1995.

Soepraptohardjo, M., and H. Suhardjo, Rice soils of Indonesia, in *Soils and Rice*, pp. 99–115, Int. Rice Res. Inst., Los Baños, Philippines, 1978.

van Bodegom, P. M., and J. C. M. Scholten, Microbial processes of methane production in a rice paddy soil: Model and experimental validation, *Geochim. Cosmochim. Acta*, *65*, 2055–2066, 2001.

van Bodegom, P. M., P. A. Leffelaar, A. J. M. Stams, and R. Wassmann, Modelling methane emissions from rice paddies: Variability, uncertainty and sensitivity analysis of processes involved, *Nutr. Cycles Agroecosyst.*, *58*, 231–248, 2000.

van Bodegom, P. M., R. Wassmann, and T. M. Metra-Corton, A process-based model for methane emission predictions from flooded rice paddies, *Global Biogeochem. Cycles*, *15*, 237–247, 2001a.

van Bodegom, P. M., P. H. Verburg, A. Stein, S. Adiningsih, and H. A. C. Denier van der Gon, Effects of interpolation and data resolution on methane emission estimates from rice paddies, *Environ. Ecol. Stat*., in press, 2001b.

van de Ven, T., and P. Tempel, ISIS 4.0, ISRIC Soil Information System: User manual, rev. ed., *Tech. Pap. 14*, Int. Soil Ref. and Inf. Cent. (ISRIC), Wageningen, Netherlands, 1994.

Verburg, P. H., A. Veldkamp, and J. Bouma, Land use changes under conditions of high population pressure: The case of Java, *Global Environ. Change*, *9*, 303–312, 1999.

Voltz, M., and R. Webster, A comparison of kriging, cubic spines and classification for predicting soil properties from sample information, *J. Soil Sci.*, *41*, 473–490, 1990.

Walter, B. P., M. Heimann, R. D. Shannon, and J. R. White, A process-based model to derive methane emissions from natural wetlands, *Geophys. Res. Lett.*, *23*, 3731–3734, 1996.

Wassmann, R., X. J. Shangguan, M. Tölg, D. X. Cheng, M. X. Wang, H. Papen, H. Rennenberg, and W. Seiler, Spatial and seasonal distribution of organic amendments affecting methane emissions from Chinese rice fields, *Biol. Fertil. Soils*, *22*, 191–195, 1996.

Whittenbury, R., K. C. Phillips, and J. F. Wilkinson, Enrichment, isolation and some properties of methane-utilizing bacteria, *J. Gen. Microbiol.*, *61*, 205–218, 1970.

P. M. van Bodegom, Free University, Faculty of Biology, Department of Systems Ecology, De Boelelaan 1087, 1081 HVAmsterdam, Netherlands. (bodegom@bio.vu.nl)

H. A. C. Denier van der Gon and P. H. Verburg, Laboratory of Soil Science and Geology (W.U.), P.O. Box 37, 6700 AA, Wageningen, Netherlands.